Dynamical Climatology

Potential vorticity in the stratosphere derived using data from satellites.

by

S.A.Clough, N.S.Grahame and A.O'Neill

DCTN 4 September 1984

Meteorological Office (Met. O. 20)
London Road
Bracknell
Berkshire RG12 2SZ

DYNAMICAL CLIMATOLOGY TECHNICAL NOTE NO. 4

POTENTIAL VORTICITY IN THE STRATOSPHERE DERIVED USING DATA FROM SATELLITES

S A CLOUGH

N S GRAHAME

AND

A O'NEILL

Submitted to Quart. J. Roy. Meteorol. Soc., July 1984

Met O 20 (Dynamical Climatology Branch)
Meteorological Office
London Road
Bracknell
Berkshire RG12 2SZ

September 1984

Note: This paper has not been published. Permission to quote from it should be obtained from the Assistant Director of the above Meteorological Office Branch.

Summary

Maps of Ertel's potential vorticity, Q, are computed on an isentropic surface in the middle of the stratosphere. They are derived using data from Stratospheric Sounding Units on board NOAA satellites. The reliability of the maps is demonstrated mainly by the conservation of Q that is shown in successive analyses, and by the close agreement obtained using independent data from two satellites.

The maps are used to follow the movement of material during the course of strong disturbances which occurred in December 1981. They clearly show the breaking of planetary waves in the stratosphere, an inherently non-linear process. The effect of wave breaking on the structure of the westerly vortex is considered.

The behaviour of diagnostics based on zonal averaging, such as the Eliassen-Palm flux, is accounted for by local changes in the distribution of Q. The implications of our findings for the study of stratospheric dynamics are discussed.

1. Introduction

Results based upon the conservation of potential vorticity are of great utility when studying rotating fluids such as the Earth's atmosphere, and form a unifying theme in dynamical meteorology (e.g. Gill, 1982). A relation requiring no assumptions about the direction of vorticity or about the ratio of horizontal to vertical scales was derived by Ertel (1942). It involves a quantity, Q, now often referred to as Ertel's potential vorticity, defined by

$$Q = \rho^{-1} (\underline{\nabla} \times \underline{u} + 2\underline{\Omega}).\underline{\nabla}\theta \qquad (1)$$

using conventional notation as detailed in the Appendix. The theorem states that if the motion is adiabatic so that

$$\frac{d\theta}{dt} = 0 \qquad (2)$$

and if it is also frictionless, then

$$\frac{dQ}{dt} = 0 \qquad (3)$$

For hydrostatic flow characteristic of atmospheric motions on the synoptic or larger scale, Eq. (1) can be simplified by writing Q as

$$Q = - g (\zeta_\theta + f) \frac{\partial\theta}{\partial p} \qquad (4)$$

where the relative vorticity ζ_θ involves differentation of the horizontal wind components on an isentropic surface (see Dutton, 1976, for a derivation). As fields of meteorological variables are generally held on isobaric surfaces, a convenient approximation to Q is

$$Q \approx - g (\zeta_p + f) \frac{\partial \theta}{\partial p} \qquad (5)$$

where derivatives are taken on isobaric surfaces. This approximation is valid for flows with large Richardson number, such as the large-scale circulations characteristic of the stratosphere (Hartmann, 1977).

Insofar as Q and θ are conserved along a trajectory, the intersection of surfaces of constant Q and θ mark the material lines of the fluid. The behaviour of the material lines can therefore be followed as the flow evolves by examining maps of Q evaluated on isentropic surfaces. The contours of Q thus presented give information of a Lagrangian nature, limited in practice by the resolution of measurement, errors in calculation and non-conservation of Q.

The advent of measurements from satellite-borne radiometers allows global monitoring of the stratospheric circulation and presents the opportunity to study aspects of the dynamics of this circulation using isentropic maps of Q. Analyses of motions based on measurements with radiosondes are less satisfactory for levels above the lower stratosphere because their precision deteriorates progressively with height. Moreover, such measurements have comparatively poor horizontal resolution and calculations of Q would be of little value where data are sparse. Analyses based on a mixture of data from radiosondes and satellites might contain spatial

inhomogeneities due to differences between instruments which would be reflected in the maps of Q. It may be advantageous therefore to use homogeneous analyses derived using data from one satellite when calculating Q, since this involves a high order of spatial differentiation of retrieved fields of thickness or temperature.

A strong gradient of Q at middle latitudes in the winter stratosphere is associated with the cold westerly vortex that dominates the circulation. The propagation of Rossby waves from the troposphere into the stratosphere depends upon this gradient, which is set up by the variation of the Coriolis parameter with latitude together with an important contribution associated with the distribution of velocity. When disturbances propagate upwards into the winter stratosphere contours of constant Q buckle, and because the amplitude of disturbances increases with height in the lower and middle stratosphere as the density decreases there is the possibility that 'wave breaking' will occur at some level, particularly where the winds are weak. This was suggested by McIntyre (1982) in his discussion of the dynamics of sudden warmings. Such breaking would be signalled by the irreversible (time orientated) deformation of the material lines of the fluid when contours of constant Q no longer just undulate back and forth but buckle beyond the point where they can unbuckle in a time-reversed way. McIntyre and Palmer (1983, and sequel 1984; both referred to hereafter as MP) presented isentropic maps of Q for the middle stratosphere which, they argued, reveal wave breaking. This was corroborated by maps of the distribution of ozone presented by Leovy et al. (1984).

MP consider irreversible mixing of Q to be an important element of wave breaking, proposing that effective mixing arises as tongues of relatively high or low Q are formed which lengthen and narrow, eventually losing their identity to instruments of limited resolution. In regions where such mixing occurs, a coarse-grain (limited resolution) measure of Q would not be conserved and gradients of Q may be permanently reduced in some places and increased elsewhere (see Fig. 5 of McIntyre, 1982). MP account in this way for the strong mid-latitude gradient and weaker equatorward gradient of Q shown on the maps they present. They refer to a 'surf-zone' or vigorously mixed region surrounding the westerly vortex of the winter stratosphere.

One way that mixing associated with wave breaking would be important for the dynamics of the stratosphere would be in permanently changing the distribution of Q, with consequent effects on the outcome of later large-scale disturbances. MP propose that erosion of the westerly vortex by mixing during wave breaking can pre-condition the stratosphere for a subsequent major warming by confining further disturbances to a smaller westerly vortex.

The role of mixing by the generation of smaller scales of motion during wave breaking remains to be quantified. In some cases the mixing may be incomplete and, moreover, the transport of Q will not just be down the existing gradient of Q. We present in this paper isentropic maps of Q for the stratosphere during wave breaking events, when the effective mixing of Q was apparently far from complete. The maps were derived using data from Stratospheric Sounding Units (SSUs) on board NOAA satellites. In

Section 2, we discuss the method used to analyse the radiometric data and the approximations used in the calculation of Q. The reliability of our maps is assessed in Section 3 by examining the extent to which Q is conserved for features reproduced on successive daily analyses, and by comparing fields derived from data obtained with another SSU on a different satellite. Data from a limited number of radiosonde/rocket ascents are also used to assess the satellite analyses, and the main features of the circulation to which we refer are confirmed by comparison with completely independent analyses obtained using data from radiosondes and rockets. The analysed fields are shown to be of sufficient quality to be valuable in diagnostic studies of the stratosphere. Maps of Q are used in Section 4 in a diagnostic study of wave breaking during a 'Canadian Warming' which occurred in December 1981. Non-linear advection on an isentropic surface is shown to have been an important element of the dynamics, and the consequences for the distribution of Q in the stratosphere are considered. It is noted that while a large area of relatively high Q was broken away from the main vortex it remained substantially unmixed and subsequently re-combined with the vortex. The overall gradients of Q were not significantly changed in the process. Section 5 contains a study of some dynamical features of the event using zonally averaged diagnostics. The evolution of such diagnostics is related to the changing distribution of Q. The conclusions are listed in Section 6.

2. Data, method of analysis and the calculation of Q

(a) Data

The data used in this study were obtained primarily from Stratospheric Sounding Units on board the satellites NOAA-6 and NOAA-7 of the Tiros-N series. The instruments were developed by the Meteorological Office; their features and performance have been described by Pick and Brownscombe (1981; hereafter PB). We will summarise the main elements of the analysis of received data emphasising those particularly relevant to the calculation of isentropic fields of Q.

The satellites are in independent, sun-synchronous orbit at a height of 850 km, providing approximately 14 orbits of data per day. The analyses are produced from radiances measured by the SSU, and also by a Microwave Sounding Unit (MSU) and a High Resolution Infra-Red Sounder (HIRS-2). An important point when assessing derived maps of Q is that all three nadir sounding radiometers scan across the orbital path, though with differing patterns and resolution. This scanning feature leads to much better horizontal resolution than a fixed angle device. The observations from the three instruments are interpolated and averaged to produce effective fields of view of about 10° half-width at 30° and 10° from nadir. The sub-orbital track reaches latitude 81° which, with scanning, gives coverage to 87° in swaths of approximately 1000 km width (comprising four independent fields of view). The pattern of orbits is such that the largest regular regions without data in the daily cycle are about 1500 km from east to west

where ascending and descending parts of orbits cross in a band between 20° and 30° from the equator, though at high latitudes the side scanning results in substantial overlap of observations. Because the orbits of NOAA-6 and NOAA-7 precess at slightly different rates, orbital tracks initially in phase move out of phase in about four days.

The weighting functions of the instrument channels used in this study are shown by PB. The three channels of the SSU have weighting functions centred in the middle stratosphere near 15,5 and 1.5 mb; the MSU channels make measurements of the lower stratosphere; and the HIRS channels have rather broad weighting functions which span much of the stratosphere. The effective vertical resolution is difficult to determine for the system. The SSU weighting functions have half-widths in the vertical of about 12-15 km, but because of the overlap of several weighting functions the effective resolution is probably about 10 km.

(b) Method of analysis

Thicknesses of the 100-20, 100-10, 100-5, 100-2 and 100-1 mb intervals are retrieved from the equivalent brightness temperatures (temperature weighted by the instrumental weighting function) by statistical regression, as described by PB. The regression coefficients are determined using thicknesses derived from a set of 1200 rocket soundings by a principal component analysis also described by PB. In order to allow for a range of variation of basic atmospheric

structure, the regression coefficients vary sinusoidally between summer and winter extremes, and are based on data from rockets grouped into seven latitudinal and seasonal zones.

The resulting profiles of thicknesses are interpolated on a 5° x 5° latitude-longitude grid using linear combinations of observations near each grid point. The weights for the interpolations are simple linear functions of distance and time, vanishing outside 5° of the grid point and 12 hours from the nominal time of the analysis. Values at points where there are no data are determined by linear interpolation from surrounding points. The fields are then smoothed by Fourier analysis with truncation at wave number 12 in the meridional (pole to pole) and zonal directions. This value is a compromise in that some information is lost at high latitudes where the orbital tracks are closely spaced. At the latitudes where ascending and descending parts of orbits cross, however, there are effectively only 14 groups of observations at a latitude so that harmonics with zonal wavenumber greater than about 7 are not well resolved.

Fields of geopotential height are produced by adding the thicknesses to an objective analysis of geopotential height at 100 mb provided by the National Meteorological Centre (Washington). NMC analyses are used for the troposphere, and also at 50 mb to supplement the vertical resolution in the lower stratosphere, although these do not affect the calculation of the maps of Q at the level presented here. Analyses of

thicknesses are independent in time in that no information about the actual state of the stratosphere is retained as a background field for subsequent analyses.

PB and Nash and Brownscombe (1983) have discussed the validation of SSU data and analyses based on them. Nash and Brownscombe primarily considered the precision of the radiances observed by the SSUs, which they evaluated by laboratory tests and by monitoring of the instrument in orbit. For our study, we have used mainly analyses derived from the SSU on NOAA-6 which Nash and Brownscombe have shown to be extremely stable over several years. The SSU on NOAA-7 was less stable over periods of a few months and there are small systematic errors in the derived analyses, which we use only for comparison.

The most important limitations in these analyses are imposed by vertical resolution, horizontal coverage and interpolation. Some indication of the effects of deficiencies in the coverage and interpolation can be gained by comparing analyses from satellites in independent orbit, while vertical resolution and retrieval problems may be studied by comparison with rocket soundings. Differences between analyses of geopotential height for NOAA-6 and NOAA-7 are typically less than 10 dam in the lower and middle stratosphere, as illustrated by PB.

Comparison of analysed geopotential heights with those calculated from radiosonde and rocket observations is carried out routinely. Typical differences at 10 mb are between about 5 and 10 dam. During

stratospheric warmings, however, differences of more than 70 dam have been found in a region of intense warming. Such large errors can arise firstly from sharp, unresolved gradients of temperature in the vertical; and secondly, but probably to a lesser extent, because standard rocket observations on which regression coefficients are based may not be statistically representative of the unusual profiles which occur during warmings. Errors will also arise when the circulation is changing quickly as observations are interpolated to a synoptic hour.

(c) The calculation of isentropic maps of Q

To avoid large errors near the pole with finite differencing in spherical co-ordinates, the isobaric height analyses are first interpolated onto a polar-stereographic grid, the grid size being 580 km at 50°N. As winds are not directly measured, they are computed using the geostrophic approximation. This approximation leads to Q being over-estimated typically by about 10% in the centre of the westerly vortex, with an error of comparable magnitude where the circulation is strongly anti-cyclonic. These estimates were obtained by comparing fields of Q computed with and without the geostrophic approximation from data provided by a multi-level numerical model based on the primitive equations (that used by Butchart et al., 1982, but with a higher resolution of 5° x 5° in the horizontal). Because the gradient wind approximation requires a higher order of spatial differentiation of the analysed height fields, we feel that the systematic error in Q resulting from the geostrophic approximation is

preferable to locally large and uncertain errors that could arise when using gradient winds. However, maps of Q are also produced routinely using gradient winds as a check.

Potential vorticity, Q, is calculated on an isentropic surface poleward of 20° latitude using a finite difference equivalent of Eq. (5). The quantities ζ_p and $(\partial\theta/\partial_p)$ are interpolated linearly with respect to log (pressure) to the chosen isentropic surface. Comparison with fields for the middle stratosphere obtained using the more accurate form, Eq. (4), shows that differences are at most about 5% in regions of strong gradients when geostrophic winds are used in both cases. No further smoothing is applied to the derived fields of Q. We present maps for the 850 K isentropic surface as this surface lies near 10 mb where the two lowest channels of the SSU overlap.

3. The reliability of isentropic maps of Q

Maps of Q on isentropic surfaces are obtained by a series of numerical processes involving the finite differentiation, interpolation and smoothing of fields. All of these can lead to systematic errors, as will approximations such as the use of geostrophic winds. Regression coefficients used in the retrieval are based on rocket measurements covering a limited range of atmospheric conditions, and the orbital patterns of the satellites are such that the spatial coverage is not uniform. Height analyses at 100 mb to which retrieved thicknesses are

added will also contain errors. Moreover, the radiometers sample over a finite volume of the atmosphere resulting in averaged or coarse-grain maps of Q.

While the SSU has the advantage of side-scanning which improves horizontal resolution, the vertical resolution is a limitation of this nadir sounding instrument. Because in the retrieval process temperature disturbances are spread in the vertical, horizontal variations in the atmosphere at different levels are averaged to some extent. It is therefore possible for features on potentially resolvable scales in the horizontal to be poorly represented by having insufficient depth in the vertical.

In view of the complexity of the retrieval method and subsequent manipulation of data and because the exact state of the real atmosphere is not known, a precise estimate of errors involved cannot be attempted. The best that can be done is to compare different types of measurements and to try to understand the differences between them, realising that differences will inevitably occur between measuring systems with different characteristics. While it would be preferable to compare maps of Q derived using fully independent measurements, we only have available at present such maps from two SSUs in independent orbit. We now consider the reliability of our maps of Q by comparing the fields derived using data from these SSUs and by noting to what extent Q is conserved in a sequence of analyses. In addition, we compare values of the quantity $(\partial\theta/\partial p)$, used in the calculation of Q, with those obtained from the few available radiosonde/rocket profiles.

Figs. 1, 2 and 3 show a sequence of maps of Q and geostrophic winds on the 850 K isentropic surface for 7, 9 and 11 December 1981, during the period on which the discussion of Sections 4 and 5 concentrates. The period was chosen not only because of its dynamical interest but also because the vertical depth of changes to the stratospheric circulation was large, so the vertical resolution of the SSU is not a major limiting factor. The top field (a) is obtained using data from the SSU on the satellite NOAA-6 (the more stable SSU), and the bottom field (b) from that on NOAA-7. The data received from both instruments on these days were almost complete. Regression coefficients are specific to each instrument to allow for differences in their characteristics. While the method of analysis is shared by both satellites, as is the field at 100 mb to which thicknesses are tied, the orbits do not generally coincide. Local features common to the two analyses are thus less likely to be peculiarities of the orbital pattern. The orbital crossings at the equator are in phase on 7 December and out of phase on 11 December (the latter meaning that an ascending orbit of one satellite occurs at the same place on the equator as a descending orbit of the other).

On all 3 days, the two sets of maps are in close agreement in both the pattern and numerical values of Q. Both show an elongated westerly vortex with local centres having high values of Q in the middle of the vortex. These centres are mostly well reproduced in each pair of analyses both in position and in numerical value, even on 11 December (Fig. 3) when the orbits are out of phase. When the coverage of both satellites is equally good, differences of up to about 10% occur in the highest values in the westerly vortex. A somewhat higher percentage difference is found in

regions where gradients of Q are weaker, such as along tongues of locally high or low Q. The tongue of relatively high Q extending westward over North America is reproduced in both sets, and local centres within the tongue are common to both. It seems unlikely that these centres are related to the orbital track of the satellite; they do not tend to lie where an orbit passes over a region and their separation does not generally correspond to the orbital spacing. Moreover, isentropic maps of Q at other levels in the lower and middle stratosphere show that they are deep features which should be resolvable by the SSU. Along the length of the tongues, the gradient of Q is small and some variation between analyses is to be expected in the position and numerical value of embedded local centres. However, an examination of many more maps for different periods shows that the break-up of such tongues into local centres is a real process in the stratosphere which can be recorded by our radiometric measurements.

The maps also show near 180°E an incursion of low Q to higher latitudes, and in Fig. 1 both analyses show that the geostrophic winds are directed across contours of constant Q at this incursion. Near the incursion, the analyses show that the westerly jet bifurcates with the strongest winds blowing over the polar cap and a weaker westerly jet at lower latitudes. Resolved gradients of Q are comparatively weak in the area separating the jets and the position and evolution of actual material lines is less certain. To ensure that the close agreement between the two sets of analyses is not merely the result of tying independently retrieved thicknesses to a common analysis at 100 mb, Q and geostrophic winds were calculated with the height of the 100 mb surface set to a constant value of

16 km. The close agreement is maintained as the example in Fig. 4 shows. The height of the 100 mb surface affects the pattern of Q and the winds mainly at low latitudes where the horizontal gradients of the 100-10 mb thicknesses are weak. Fig. 4 also confirms that the local centres of high Q in the western hemisphere, shown in Fig. 2, are not imposed hydrostatically by the analysis at 100 mb.

The conservation of Q provides an important objective check on the reliability of the coarse-grain maps, particularly since no constraint of continuity in time is imposed in the analysis. Even localised features can be identified throughout the period shown on both sets of analyses, e.g. the local maximum near 165°W, 40°N. The maps also show good temporal continuity in the highest and lowest values of Q throughout late November and December.

As a further test of conservation we have computed the total area where Q is greater than a specified value for several values of Q. The results are shown in Fig. 5, in which curves are plotted for one month using data from NOAA-6 with one curve duplicated for comparison using data from NOAA-7. If plotted contours of constant Q actually represent material lines and inviscid, adiabatic conditions can be reasonably assumed, then the integrated areas should be nearly constant for quasi non-divergent flow. The areas are not fixed, however, but vary systematically (and not monotonically) on a time-scale of the order of a week. It seems that this variation is related to the limited resolution of the measurements and the lack of conservation this entails. For example, the change in the area inside the contour labelled '4' in mid December can be linked in Fig. 11

(to be discussed later) with the uncertain position of the contour where gradients of Q are weak over Canada. The curves, however, show very little random variability; even when the flow became highly contorted in early December, areas are conserved remarkably well. Apart from a systematic difference (due to instrumental changes uncorrected in the analysis of data from NOAA-7), areas computed with data from the two satellites are in close agreement in Fig. 5, even though the orbital patterns went in and out of phase several times during the period shown.

As a preliminary assessment of the limitation imposed by vertical resolution, we have computed values of $(\partial\theta/\partial_p)$ in Eq. (5) at various levels in the stratosphere using data from all available combined radio-sonde/rocket ascents (14 in total) during the period late November to mid December. The values of $(\partial\theta/\partial_p)$ were compared with those derived from co-located vertical profiles retrieved from measurements by the SSU. At the 850 K isentropic surface near 10 mb, differences were up to about 10%. We caution, however, that this applies to a particular period and limited set of rockets; this value can be a significant underestimate of differences where vertical structure is only partially resolved by the SSU, such as may arise during strong warmings. As discussed later, the stratospheric disturbances in this period should be resolvable as they showed comparatively little westward tilt with height and extended over a large depth.

We conclude from the conservation of Q, temporal continuity of patterns and their reproducibility using different data that, for our period of study, maps of Q can be used to follow the movement of material in the stratosphere, albeit with a coarse-grain view.

4. The dynamics of a strong stratospheric disturbance in early winter

During November - December 1981, a strong disturbance occurred in the stratosphere with the warmest air being drawn north eastwards over Canada as the predominantly westerly flow became meridional in places. Fig. 6 shows the brightness temperature measured by the lowest channel of the SSU (centred near 15 mb) for 25 November and 7 December. While the temperature pattern became highly distorted, the temperature of the warm pool actually decreased, and changes both locally and in the zonal mean were smaller than those observed during some stratospheric warmings later in winter. Naujokat et al. (1982) have noted, however, that the warming showed the most intense development of height in zonal wavenumber 1 observed in the Northern Hemisphere since 1965.

Warmings of this type have been called Canadian warmings by Labitzke (1977). They tend to occur in early winter and to be confined to the lower and middle stratosphere. For the case in 1981, the increase in temperature over the pole was greater at 30 mb than at 10 mb, and the top channel of the SSU (centred near 1.5 mb) did not detect much warming. However, the vertical depth of the changes in temperature was such that, at 10 mb in early December, easterly winds replaced westerlies in the zonal mean from the pole to almost 60°N. Based on the zonal winds, the warming could

almost be called 'major' according to the definition used by the World Meteorological Organisation, although it may be contrasted with major warmings occurring later in winter (such as that described by Palmer, 1981) in that temperature changes were comparatively small and the westerly vortex was not completely split into well separated centres of circulation. In terms of the movement of Q, however, the disturbance is a clear example of wave breaking in the stratosphere in that it involved the irreversible deformation of material lines over great distances.

We now discuss the dynamics of the warming by first examining a sequence of maps of Q and geostrophic winds evaluated on the 850 K isentropic surface. For deep perturbations in the lower and middle stratosphere, such as we study here, Fels (1982) has calculated that the radiative damping time is on the order of 20 days. This justifies taking Q contours as material lines for an event which evolves over a week or so.

The build-up of the disturbance in late November was characterised by air with low values of Q being drawn north eastwards around the westerly vortex, apparently cutting through the outer edge of the vortex. This is shown in Figs. 7 and 8 for 25 and 30 November. As the area of low Q reached polar latitudes, a tight gradient of Q was established to the north in a region of cross-polar flow. Fig. 9 shows that by 4 December, the main westerly vortex was markedly elongated, extending into a tongue of high Q to the south of the areas of low Q. By 7 December, Fig. 1a shows that an area of low Q has been drawn right across polar regions to pinch off the tongue of high Q from the main vortex near 0°E by 9 December (Fig. 2a). Q is a dynamically active quantity since its evolving distribution feeds back

on the wind field which advects it. This is strikingly illustrated in Fig. 1a by the giant meander formed in the flow around adjacent tongues of high and low Q. A comparison of Fig. 1a with Fig. 7 shows that the strong cross-polar jet has rotated anti-clockwise as seen on the maps.

The changing wind pattern and the extreme buckling of contours clearly cannot be described as a linear perturbation about any basic state, zonally symmetric or not, at least for any length of time. Rather, quasi-horizontal, non-linear advection of Q on an isentropic surface appears central to the dynamics, the non-linearity arising because of the role Q plays in governing the advecting wind field. The non-linear nature of the process has been demonstrated by O'Neill and Pope (work to be reported) using the numerical model of Butchart et al. (1982). Changing the sign of the perturbation applied at the lower boundary of the model does not change the sign of the response in the stratosphere in the region where tongues of high and low Q are formed.

A comparison of the brightness temperatures in Fig. 6 with the corresponding maps of Q in Figs. 7 and 1a shows that the eastward and poleward moving warm centre can be related to advection of air. We suggest that non-linear advection can account for the eastward movement of thermal systems often observed during the course of stratospheric warmings, rather than the excitation of an eastward-travelling linear wave in the atmosphere.

The pattern of adjacent tongues of high and low Q in Fig. 1a is similar to the distribution of Q associated with some examples of blocking in the troposphere (G.J. Shutts, personal communication), albeit that the pattern will be typically on a larger scale in the stratosphere. Fig. 1a shows that the jet bifurcated near 180°E, with one branch becoming meridional and the other remaining westerly. The two streams re-combined near 30°E. The divided current enclosed cyclonic and anti-cyclonic regions, which is a characteristic feature of one type of blocking development in the troposphere (see, for example, Namias and Clapp, 1951, Fig. 11 and associated discussion). Despite the similarity of the patterns, however, there may be important dynamical differences to be identified in the evolution of the stratospheric and tropospheric systems.

Our sequence of maps shows that as the long tongue of high Q became separated from the main westerly vortex, distinct local centres of high Q formed along its length. These centres can be located on successive days on our maps, three being shown clearly in Fig. 2a for 9 December. They seem to develop near the exit of the strong jet which forms north of the incursion of low Q. A series of them appear as the axis of this jet rotates, which suggests that they may be formed when the tongue of high Q is broken up by comparatively small disturbances steered downstream along the jet. MP have also noted that the necessary condition for instability (Charney and Stern, 1962) is satisfied locally since the meridional gradient of Q in the tongue changes sign, and they propose that, as a result of this instability, local centres of high Q emerge where filaments of material are entrained in a small region. They argue that this would make the mixing of Q incomplete on a coarse-grain map.

The circulation shown in Fig. 2a is in good agreement with that shown in Fig. 10, a corresponding and independent analysis of geopotential height at 10 mb kindly prepared for us at the Free University of Berlin. It is a hand-drawn analysis of data from radiosondes and rockets. The direct wind observations confirm the presence of a closed cyclonic circulation over North America - a 'cut-off low', reminiscent of those formed during blocking in the troposphere. The other centres of high Q on our maps lie over oceans where there are very few radiosonde observations.

Fig. 3a for 11 December shows that the main vortex was elongated into two distinct centres of high Q. (The corresponding Berlin analysis confirms the two cyclonic centres at 30 mb, but in the analysis at 10 mb there are not enough observations to resolve two centres.) An elongated vortex is a common feature of stratospheric warmings, whether they are labelled as wave-1 or wave-2 warmings because of the position of the vortex with respect to the pole. For this Canadian warming, it seems that the wave-2 component in the stratosphere was a concomitant of non-linear advection during wave breaking, and not just a linear response to the observed weak wave-2 component of the tropospheric flow. Non-linear advection during wave breaking may account for the observation by Labitzke (1977) that stratospheric warmings commonly involve the growth of wave 1 in the stratosphere followed by the growth of wave 2 as wave 1 declines. An initially displaced polar vortex being stretched over the pole in the course of wave breaking is consistent with this evolution.

On our coarse grain maps, the local centres of high Q had partially re-combined with the main vortex when further wave breaking occurred. Fig. 11 for 18 December shows an early stage of a development which proceeded in a similar way to that in early December; low Q was drawn around the westerly vortex across the polar cap and a tongue of high Q was temporarily separated from the main vortex. (The increase in the shaded area in Fig. 11 compared with earlier figures is likely to be the result of limited resolution, as noted in Section 3.) Such a succession of events through winter appears as pulsations in the strength of the stratospheric Aleutian High. Injections of low Q from equatorial latitudes during wave breaking must maintain the Aleutian High against dissipation as a feature in a time-mean chart. Shutts (1983) has invoked a similar argument for the maintenance of blocking anti-cyclones in the troposphere.

Despite the disruption of the vortex by strong perturbations in quick succession, Fig. 12 shows that on 25 December the vortex was similar to its state one month earlier (Fig. 7). This is confirmed by Fig. 5 in that areas contained within contours of fixed Q are about the same at the beginning and end of the period. The westerly vortex on these occasions was comparatively undisturbed so there is less uncertainty in its structure because of limited resolution of measurements. Any mixing effect of wave breaking on a coarse-grain distribution of Q was clearly limited in our examples, or was offset by the non-conservative effect of radiation over the period. An important task for the future would be to assess the role of the latter.

5. Diagnostics based on zonal averaging

Having recorded that advection of Q on isentropic surfaces was a principal dynamical feature of the Canadian warming of December 1981, we now consider how this advection is reflected in the behaviour of zonally averaged diagnostics, such as Eliassen-Palm fluxes introduced by Andrews and McIntyre (1978).

Fig. 13 for 30 November shows the direction and divergence of the total Eliassen-Palm (EP) flux comprising all resolved harmonics around latitude circles. The divergence is the quantity denoted as D_P (the effective zonal force per unit mass) by Edmon et al. (1980). The EP fluxes lean distinctly equatorwards in the stratosphere at low latitudes, but are more vertically aligned at high latitudes. In the corresponding map of Q (Fig. 8), troughs and ridges in the flow formed by adjacent tongues of low and high Q have a south-west to north-east orientation, corresponding to a poleward eddy momentum flux and hence an equatorward EP flux in Fig. 13. This may be termed the 'absorbing' stage of wave breaking in that EP fluxes point strongly to low latitudes and converge there (cf. the discussion of critical layers by McIntyre, 1982). At high latitudes where the flow is almost meridional across the polar cap in Fig. 8, the inflow and outflow across latitude circles are nearly off-setting in the zonally averaged eddy momentum flux and the EP flux is more vertically aligned. The strong convergence of the EP flux at high latitudes in Fig. 13 can be associated in Fig. 8 with the replacement over the polar cap of high Q by low Q,

recalling that convergence of the EP flux corresponds to an equatorward flux of potential vorticity (e.g. Edmon et al., 1980, Eq. 2.5) in the quasigeostrophic approximation.

A westward tilt with height is associated with the vertical component of the EP flux and vertical propagation of disturbances. This tilt virtually disappeared in the lower and middle stratosphere during early December as low Q continued to be drawn eastwards around the vortex. The upper stratosphere became cut off by the breaking of waves below it, and perturbation amplitudes rapidly decreased with height above 5 mb.

Fig. 14 for 9 December shows that arrows representing the EP flux lie almost horizontally towards the pole in the stratosphere, and there is strong convergence at high latitudes and divergence at lower latitudes. A few days later, the direction of these arrows is almost completely reversed. The distribution of winds for 9 December (Fig. 2a) puts in proper perspective what appears to be a notable feature of the EP fluxes. The flow is cross polar and a slight change to the curvature of the flow over the polar cap is magnified in the computation of departures from the zonal mean. Such a switching in the direction of polar EP fluxes is therefore not unexpected during major warmings as a consequence of the displacement of the westerly vortex from the pole, and has been reported by Palmer (1981) and by O'Neill and Youngblut (1982). It would not seem valid to account for the phenomenon by linear propagation of waves on the current zonal mean state of the stratosphere.

A change in the direction of EP fluxes on account of slight distortions to a cross-polar flow should be contrasted with the polar 'focusing' of the fluxes which can occur at an advanced stage of wave breaking as contours of Q become more buckled. The latter effect is most clearly seen in the idealised numerical experiment of Dunkerton et al. (1981) where, by the symmetry imposed by forcing in wave 2 only, no cross-polar flow can occur. The focusing described in their paper involves non-linear isentropic advection of Q rather than linear refraction of waves. In the atmosphere during the period of the present study, Q was continually being advected around the westerly vortex, and this 'reflecting stage' of wave breaking (see McIntyre, 1982) when a single event reaches maturity was never seen in the pattern of EP fluxes.

Cross-polar flow when the westerly vortex is displaced from the pole can lead to rapid changes in zonal mean diagnostics if, during wave breaking, low Q is advected around the vortex from low latitudes directly into the polar cap. This may occur if an initially displaced vortex (the normal situation in the winter stratosphere of the Northern Hemisphere) is subjected to a further perturbation from the troposphere. Labitzke (1977) has suggested that a pre-cursor for a major warming (defined by polar easterlies in the zonal mean) is that amplitudes of wave 1 in the stratosphere are large and those of wave 2 small, which is consistent with the occurrence of cross-polar flow. Palmer and Hsu (1983) have also pointed out that zonal mean diagnostics are sensitive to the location of the main vortex, though they considered harmonic waves propagating linearly on a displaced vortex rather than advection around the vortex during wave breaking.

The movement of the zonal mean jet to high latitudes has also been regarded as a pre-condition for a major warming (e.g. Kanzawa, 1982). This can be reconciled with Labitzke's suggestion if the zonal mean westerly jet at high latitudes does not occur in an axially symmetric flow but is the adjunct of a westerly vortex displaced from the pole (large wave 1). For example, the zonal mean winds for the flow shown in Fig. 11 are strongest near 70°N. Rather than confinement of waves by linear propagation on the zonal mean state, advection of low values of Q directly into the polar cap seems the likely reason rapid zonal mean changes are favoured.

Non-linear, local processes during wave breaking result in changes in zonal mean and eddy quantities. When the flow is highly non linear, it can be misleading to regard eddies as forcing changes in the current zonal mean if the implication is of causality. Physically, the interpretation that the eddies represent a forcing is warranted only if they evolve independently of the zonal mean, at least to some approximation. The notions of wave-mean flow or wave-wave interaction have limited value when the waves and the mean are defined with respect to a co-ordinate system that ceases to be relevant to the flow that obtains.

6. Conclusions

Difficulties for the interpretation of isentropic maps of Q derived from satellite data are the limited resolution of measurements and the non-conservative effects of radiation. Both can pose problems for the study of stratospheric warmings, particularly where large vertical gradients of

temperature arise. Our study has concentrated on an event in the middle stratosphere for which these difficulties do not appear to have been serious. We have used SSU data to calculate maps of Q and demonstrated for our case that the derived fields show good temporal continuity, conserve Q well when resolution is not a limitation, and are in close agreement when obtained using data from different SSUs in independent orbit.

In agreement with the findings of McIntyre and Palmer (1983, 1984) maps of Q show that wave breaking occurs in the stratosphere. This happens when material lines buckle beyond the point where they can merely unbuckle in a time-reversed way. The Canadian warming of December 1981 was one of the clearest examples of wave breaking we have found in an examination of 5 years of SSU data, although associated temperature changes were not remarkable. During the event an area of low Q was advected over large distances around the westerly vortex and into the polar cap as regions of high Q were temporarily separated from the vortex.

The term 'wave breaking' unifies what is observed in the stratosphere with other examples of wave breaking in different fluids, although the outcome of the phenomenon may vary with the circumstances. The breaking of ocean waves on a beach causes a substantial convergence of momentum flux as the reflected wave is weak. Our examples of wave breaking in the stratosphere appear to have behaved differently. In two such events, regions of high Q were separated from the main westerly vortex only to re-combine with it as the stratosphere became temporarily less disturbed, so that the structure of the vortex was not greatly changed in a coarse-grain map. Wave breaking in these cases did not lead, as complete mixing would, to irreversible

down-gradient transport of Q. Because the mixing of Q was limited, the stratospheric waves were, in effect, more 'reflected' from the region where they broke than ocean waves would be in striking a beach, in that the transport of Q was not one way. There is likely to be, however, a cumulative effect on the coarse-grain distribution of Q resulting from the generation of small scales of motion by a succession of breaking waves.

Wave breaking in the stratosphere is an important dynamical process because it appears, from a preliminary study we have made of several winters, to be happening nearly all of the time in the winter stratosphere of both hemispheres. In the Northern Hemisphere, the most pronounced breaking occurs in a preferred geographical location, and a monthly-mean map of Q in mid winter can show distinct tongues of high and low Q in the vicinity of the Aleutian High (V.D. Pope, personal communication).

A range of behaviour can be expected for these breaking waves. The Canadian warming of our study involved a huge breaking wave which affected a deep layer in the stratosphere and for which Q was well conserved. For other wave breaking events, coarse-grain maps of Q may show poor conservation even over a few days when vertical structure is generated below a scale that can be adequately resolved. This can happen if tongues of high and low Q are advected at different rates on different isentropic surfaces, thus sharpening vertical gradients. The vertical averaging inherent in the measurements will change the isentropic gradients of Q, in addition to any mixing on isentropic surfaces. An example where

differential advection led to substantial loss of resolved Q is discussed by Fairlie and O'Neill (to be reported) for a strong stratospheric warming in January 1982.

The stratification introduced by differential advection would be dynamically important for the subsequent evolution of the circulation. Firstly, upward-propagating Rossby waves have large horizontal and vertical wavelengths, and we suppose that their propagation is influenced mainly by the coarse-grain distribution of Q. Where Q is taken up in comparatively small vertical scales, the distribution of Q that the waves "feel" will be altered. Secondly, non-conservative changes in Q due to radiation will become more important as relaxation times are shorter where stronger vertical gradients in temperature are introduced (Fels, 1982).

We noted that wave breaking is not within the compass of linear theory, and we considered how the associated isentropic advection of Q was reflected in zonal mean diagnostics. We ascribed some well-known features of these diagnostics and their evolution to local processes, and remarked that non-linear advection of Q around a displaced polar vortex can lead to rapid changes in zonal mean quanities.

We conclude that an examination of maps of Q will be more illuminating as a test of dynamical hypotheses than a study of zonal mean quantities alone.

Acknowledgements

We are grateful to many colleagues for practical help and useful discussions. All members of the Middle Atmosphere Group at the Meteorological Office contributed to this work; in particular, Kelvyn Robertson prepared almost all the material for the figures shown, and Victoria Pope constructively discussed most aspects of the paper. Professor K Labitzke and co-workers at the Free University of Berlin very kindly prepared special analyses for several days of interest to us. We benefitted from discussions with the following:

D.G. Andrews, S.B. Fels, B.J. Hoskins, M.E. McIntyre, T.N. Palmer, and G.J. Shutts. We thank D.G. Andrews and M.E. McIntyre for their perceptive comments on the paper.

Appendix

Symbols used in the text.

f	Coriolis parameter
g	Acceleration due to gravity
p	Pressure
Q	Ertel's potential vorticity
t	Time
$\underline{u}$	Velocity
ζ	Vertical component of relative vorticity
θ	Potential temperature
ρ	Density
$\underline{\Omega}$	Angular velocity of the Earth

References

Andrews, D.G. and McIntyre, M.E. 1978 Generalised Eliassen-Palm and Charney-Drazin theorems for waves on axisymmetric mean flows in compressible atmospheres, J. Atmos. Sci., 33, 2031-2048.

Butchart, N., Clough, S.A. Palmer, T.N. and Trevelyan, P.J. 1982 Simulations of an observed sudden warming with quasigeostrophic refractive index as a model diagnostic, Quart. J. Roy. Met. Soc., 108, 475-502.

Charney, J.G. and Stern, M.E. 1962 On the instability of internal baroclinic jets in a rotating atmosphere, J. Atmos. Sci., 19, 159-172.

Dunkerton, T., Hsu, C.-P.F. and McIntyre, M.E. 1981 Some Eulerian and Lagrangian diagnostics for a model stratospheric warming, J. Atmos. Sci., 38, 819-843.

Dutton, J.A. 1976 The ceaseless wind. An introduction to the theory of atmospheric motion, New York, McGraw Hill, 579 pp.

Edmon, H.J., Hoskins, B.J. and McIntyre, M.E. 1980 Eliassen-Palm cross-sections for the troposphere, J. Atmos. Sci., 37, 2600-2616. (See also Corrigendum, ibid., 38, 1115, 1981).

Fels, S.B. 1982 A parameterization of scale dependent radiative damping rates in the middle atmosphere, J. Atmos. Sci., 39, 1141-1152.

Hartmann, D.L. 1977 On potential vorticity and transport in the stratosphere, J. Atmos. Sci., 34, 968-77.

Kanzawa, H. 1982 Eliassen-Palm flux diagnostics and the effect of the mean wind on planetary wave propagation for an observed sudden stratospheric warming, J. Met. Soc. Japan, 60, 1063-1073.

Labitzke, K. 1977 Interannual variability of the winter stratosphere in the northern hemisphere, Mon. Wea. Rev., 105, 762-70.

Leovy, C.B., Sun, C.-R., Hitchman, M.W., Remsberg, E.E., Russell, J.M., Gordby, L.L., Gille, J.C. and Lyjak, L.V. 1984 Transport of ozone in the middle stratosphere: evidence for planetary wave breaking, J. Atmos. Sci., in press.

McIntyre, M.E. 1982 How well do we understand the dynamics of stratospheric warmings? J. Met. Soc. Japan, 60, 37-65.

McIntyre, M.E. and Palmer, T.N. 1983 Breaking planetary waves in the stratosphere, Nature, 305, 593-600.

1984 The 'surf zone' in the stratosphere, J. Atmos. Terr. Phys., in press.

Namias, J. and Clapp, P.F. 1951 Observational studies of general circulation patterns. Compendium of Meteorology, American Meteorological Society, 551-567.

Nash, J. and Brownscombe, J.L. 1983 Validation of the stratospheric sounding unit, Int. Counc. Sci. Unions, Comm. Space Res., Adv. Space Res., 2, No. 6, 59-62.

Naujokat, B., Petzoldt, K., Labitzke, K. and Lenschow, R. 1982 The fourth winter of PMP-1, 1981/82: A winter with several interesting features, Beilage zur Berliner Wetterkarte, Institute fur Meteorologie, Free University of Berlin, Berlin.

O'Neill, A. and Youngblut, C.E. 1982 Stratospheric warmings diagnosed using the transformed Eulerian-mean equations and the effect of the mean state on wave propagation, J. Atmos. Sci., 39, 1370-1386.

Palmer, T.N. 1981 Diagnostic study of a wavenumber -2 stratospheric sudden warming in a transformed Eulerian-mean formalism, J. Atmos. Sci., 38, 844-855.

Palmer, T.N. and Hsu, C.-P.F. 1983 Stratospheric sudden coolings and the role of nonlinear wave interactions in preconditioning the circumpolar flow, J. Atmos. Sci., 40, 909-928.

Pick, D.R. and Brownscombe, J.L. 1981 Early results based on the stratospheric channels of TOVS on the TIROS-N series of operational satellites, Int. Counc. Sci. Unions, Comm. Space Res., Adv. Space Res., 1, No. 4, 247-260.

Shutts, G.J. 1983 The propagation of eddies in diffluent jetstreams: eddy forcing of 'blocking' flow fields, Quart. J. Roy. Met. Soc., 109, 737-761.

Figure Captions

Fig. 1: Ertel's potential vorticity, Q, and geostrophic winds evaluated on the 850 K isentropic surface for 7 December 1981; (a) using data from NOAA-6 and (b) using data from NOAA-7. Units: $Km^2kg^{-1}s^{-1} \times 10^{-4}$. In these units, areas with values of Q between 4 and 6 are shaded.

Fig. 2: As Fig. 1 for 9 December 1981.

Fig. 3: As Fig. 1 for 11 December 1981.

Fig. 4: As Fig. 2 but with analysed thicknesses tied to a 100 mb surface with a constant height of 16 km.

Fig. 5: Area ($km^2 \times 10^6$) on 850 K isentropic maps with values of Q greater than specified value. Curves are labelled in units of $Km^2\ Kg^{-1}s^{-1} \times 10^{-4}$. Solid lines using data from NOAA-6; dashed line using data from NOAA-7.

Fig. 6: Brightness temperature (K) from the lowest channel of the SSU on NOAA-6, which has a weighting function centred near 15 mb (a) on 25 November 1981; (b) on 7 December 1981.

Fig. 7: Ertel's potential vorticity, Q, and geostrophic winds evaluated on the 850 K isentropic surface for 25 November 1981, using data from NOAA-6. Units: $Km^2kg^{-1}s^{-1} \times 10^{-4}$.

Fig. 8: As Fig. 7 for 30 November 1981.

Fig. 9: As Fig. 7 for 4 December 1981.

Fig. 10: Winds and analysed geopotential height at 10 mb for 9 December 1981, provided by the Free University of Berlin.

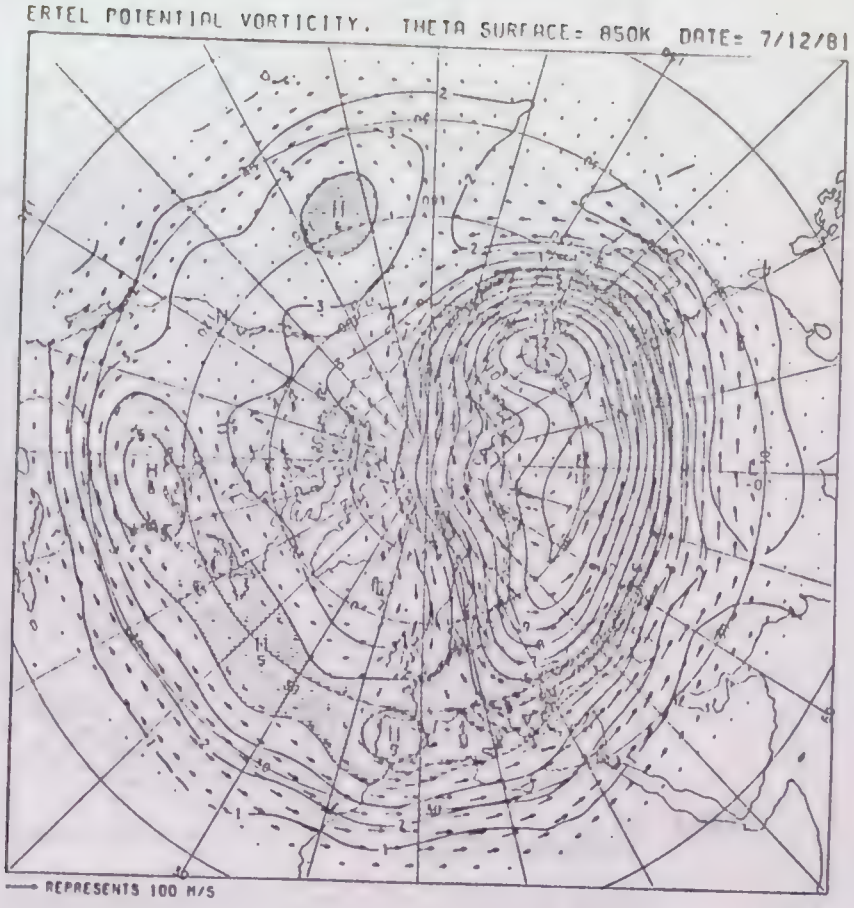
Fig. 11: As Fig. 7 for 18 December 1981.

Fig. 12: As Fig. 7 for 25 December 1981.

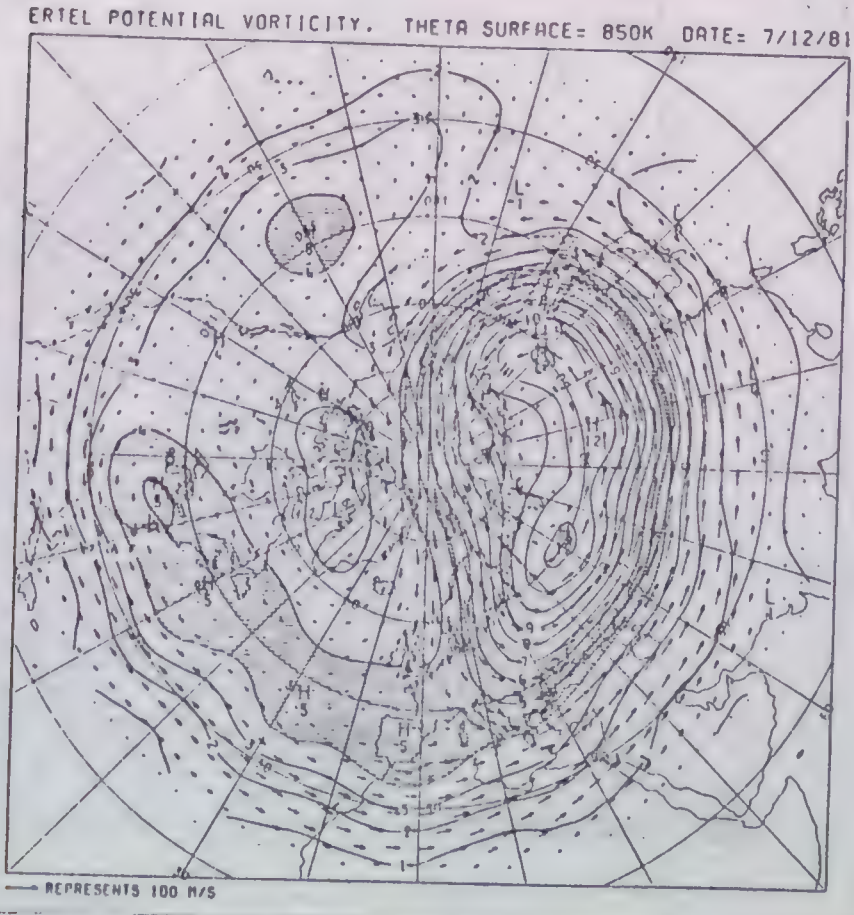
Fig. 13: Direction and divergence (see text) of Eliassen-Palm flux for 30 November 1981. Units: $ms^{-2} \times 10^{-5}$.

Fig. 14: As Fig. 13 for 9 December 1981.

Fig 1

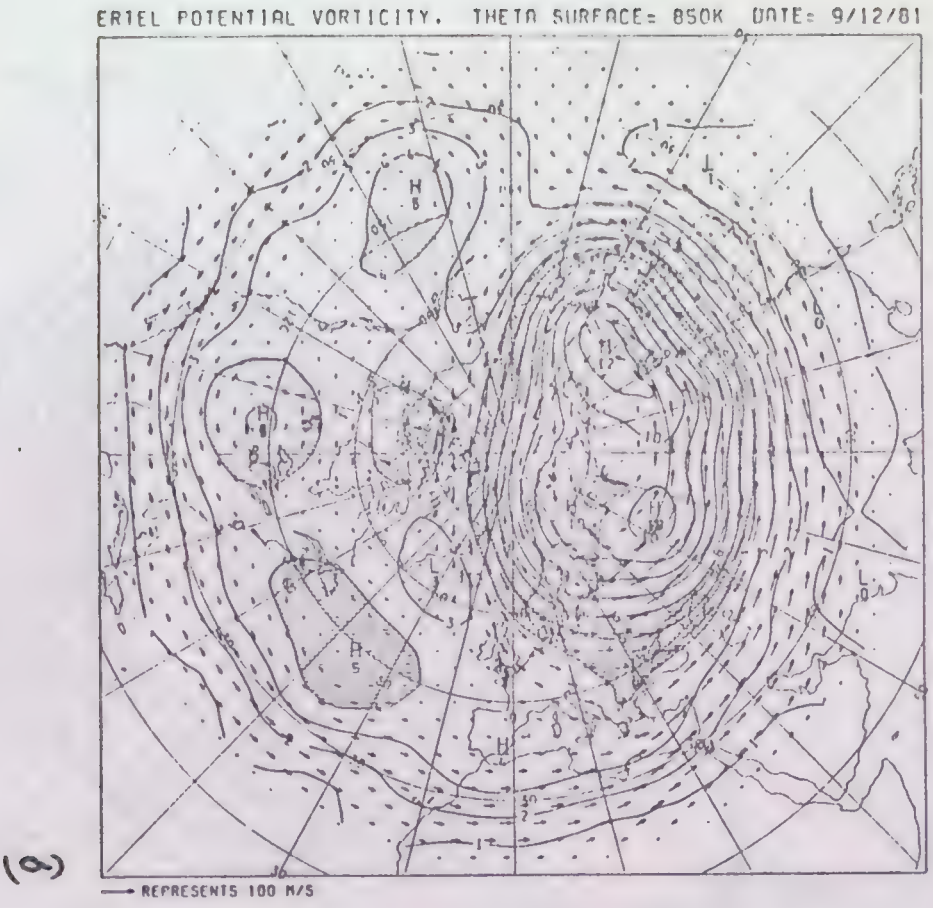

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 7/12/81
REPRESENTS 100 M/S

(a)

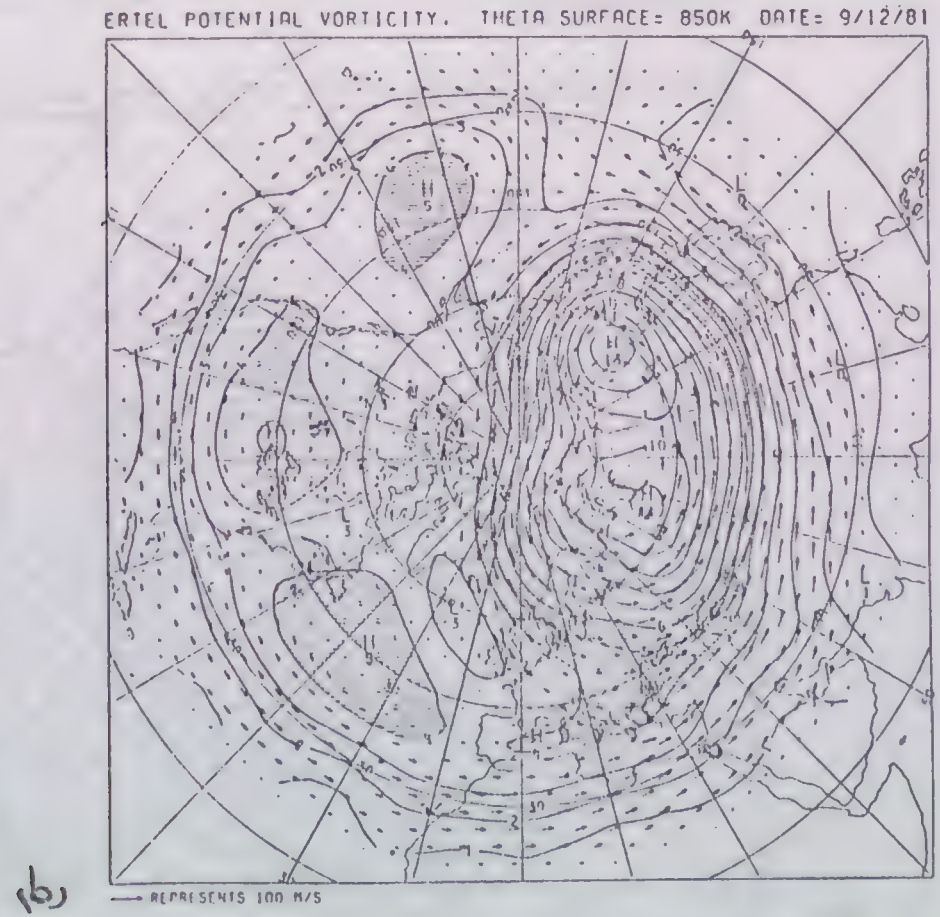

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 7/12/81
REPRESENTS 100 M/S

(b)

Fig 2

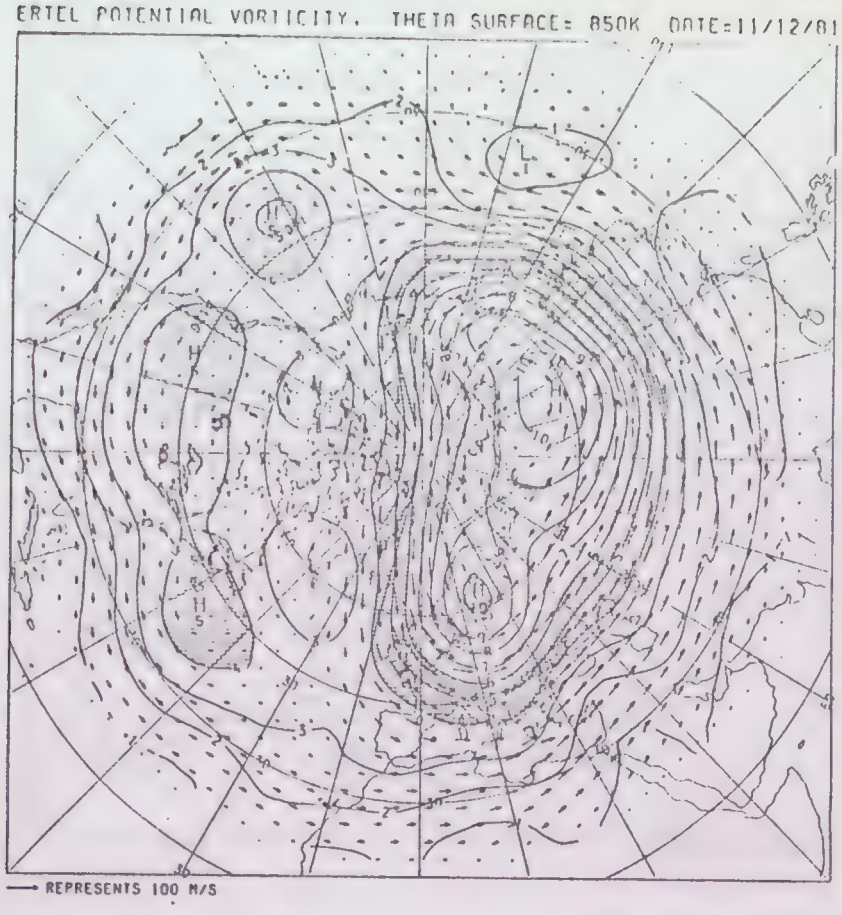

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 9/12/81
REPRESENTS 100 M/S

(a)

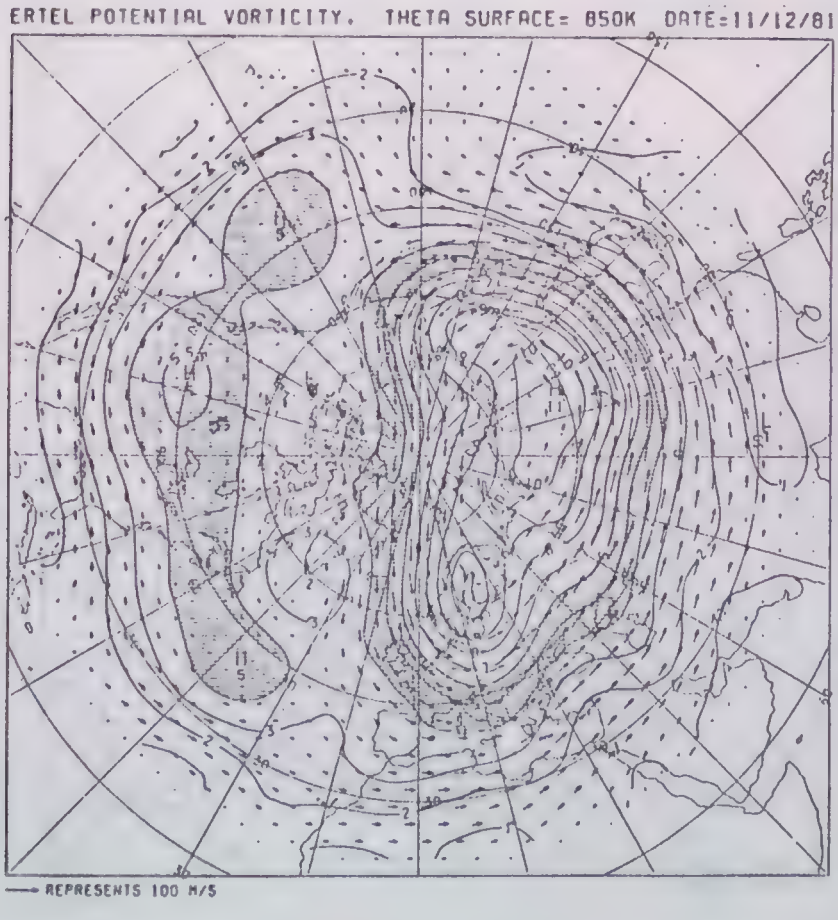

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 9/12/81
REPRESENTS 100 M/S

(b)

Fig 3

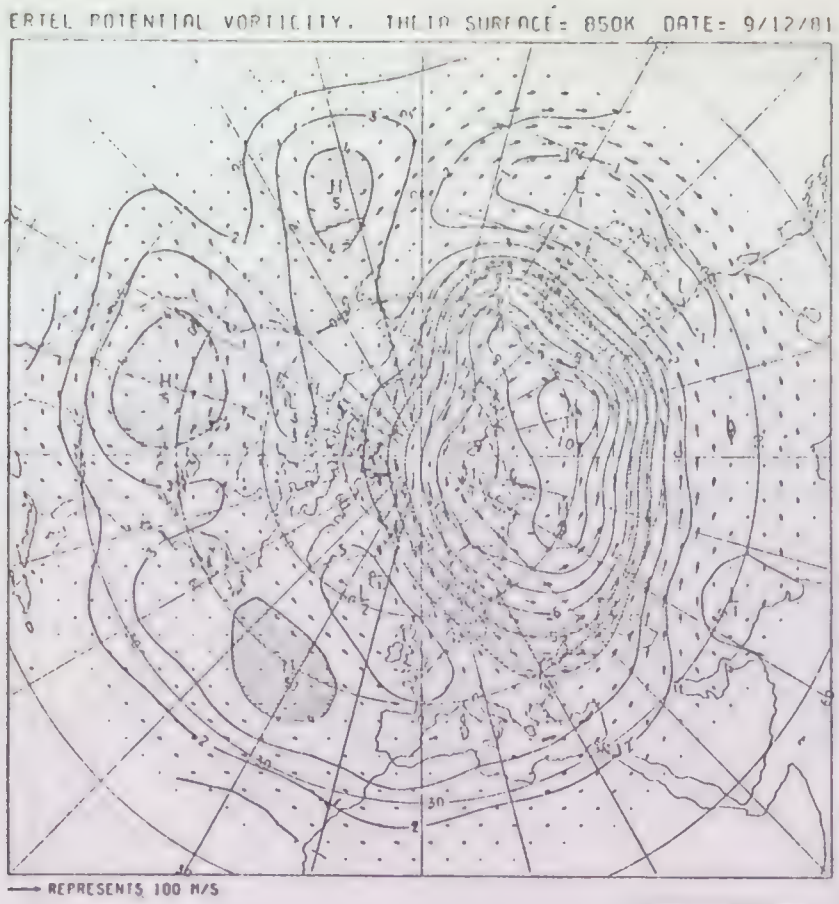

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE=11/12/81
REPRESENTS 100 M/S

(a)

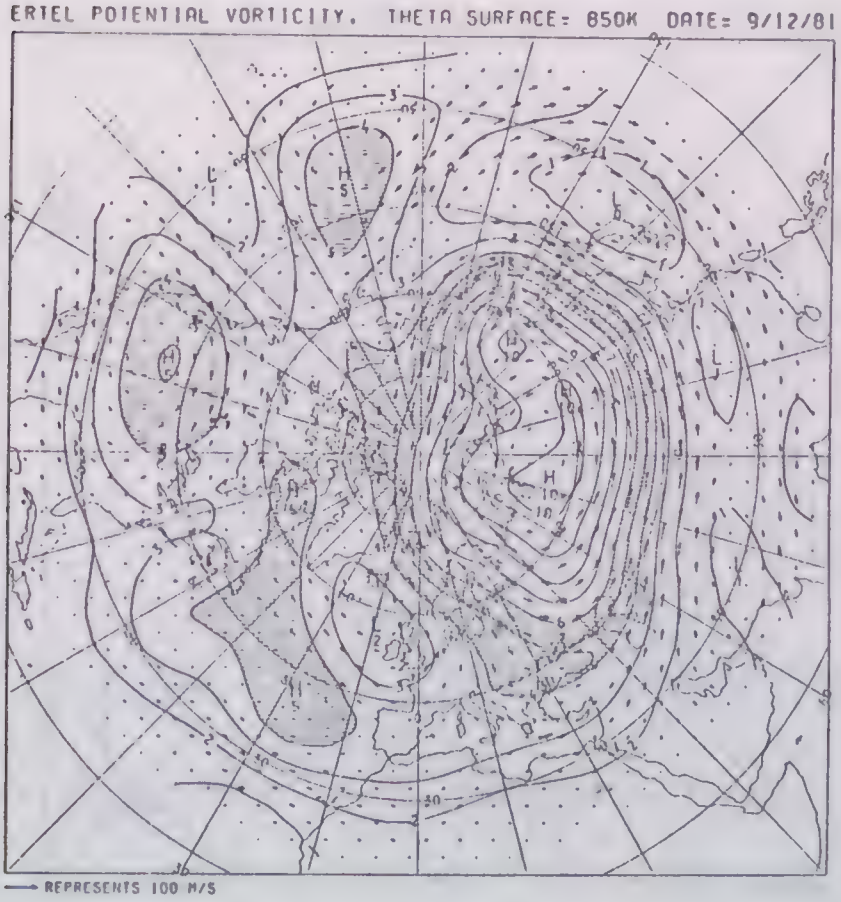

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE=11/12/81
REPRESENTS 100 M/S

(b)

Fig 4


ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 9/12/81
REPRESENTS 100 M/S

(a)


ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 9/12/81
REPRESENTS 100 M/S

(b)

Fig 5

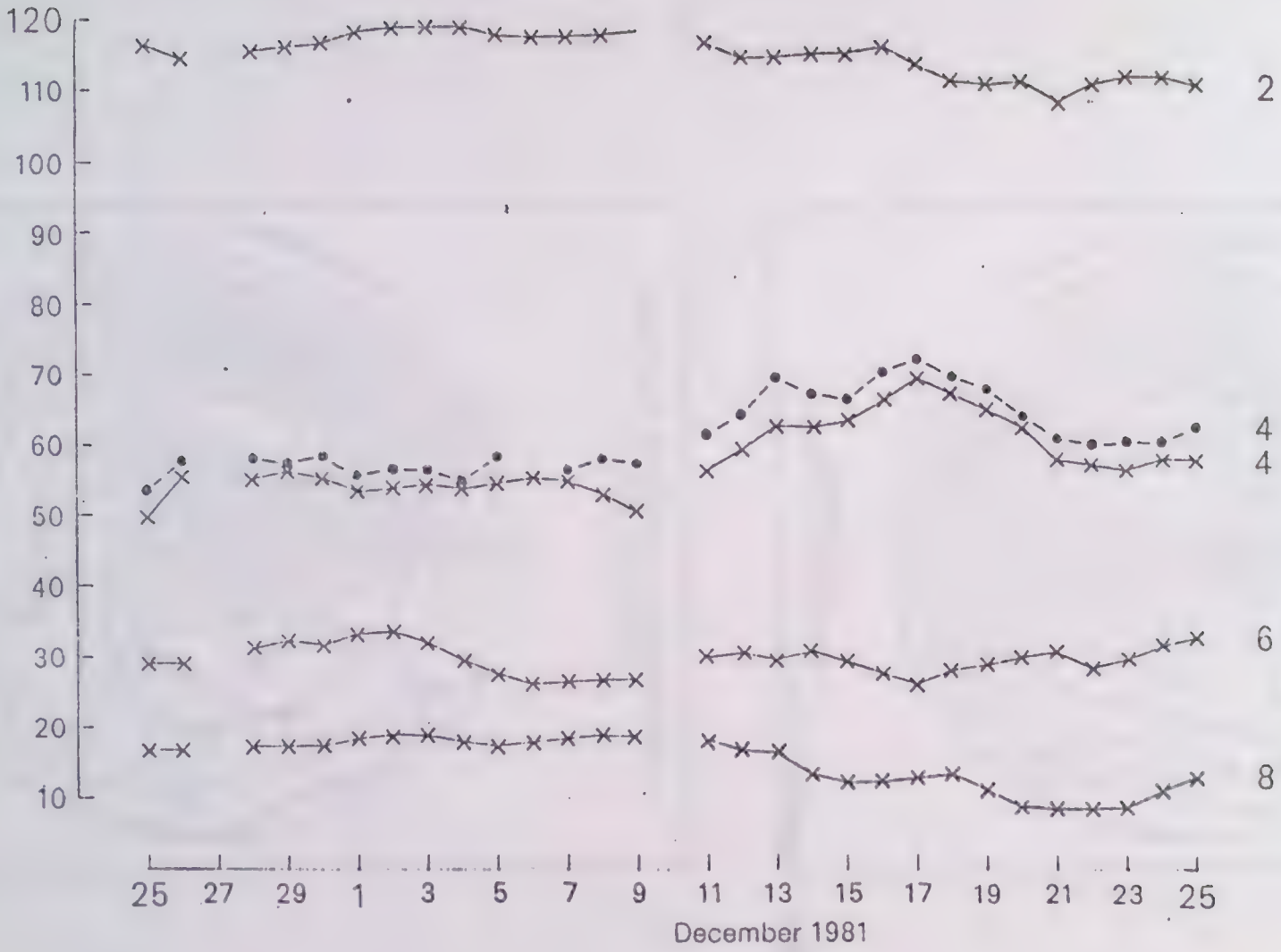
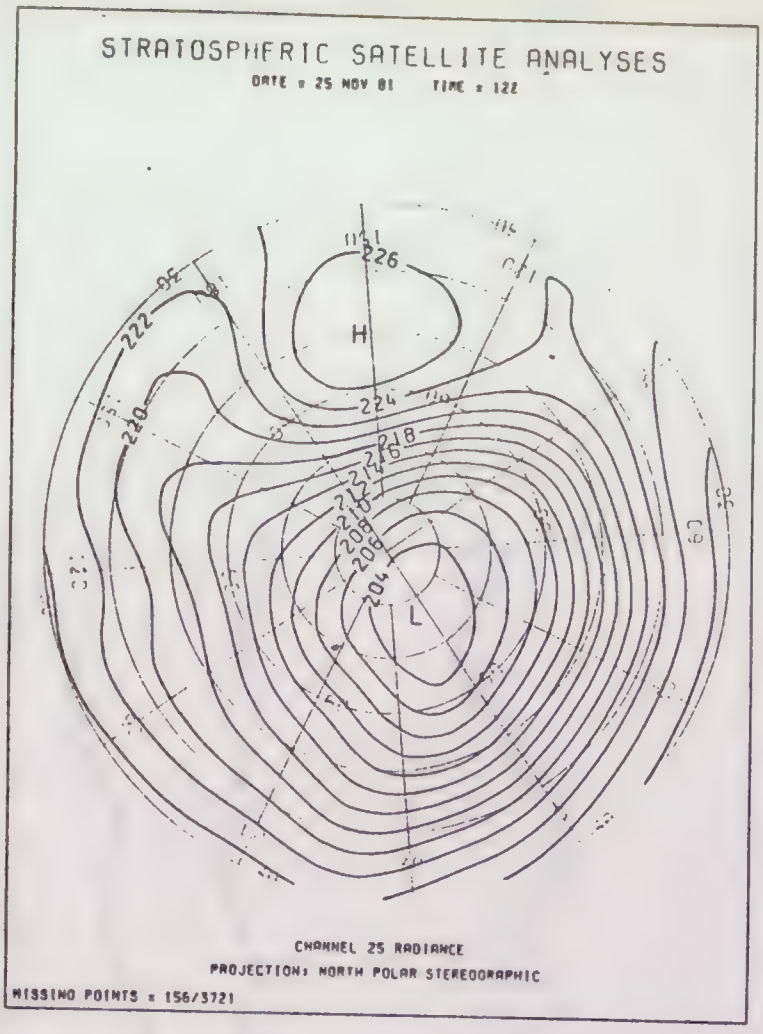
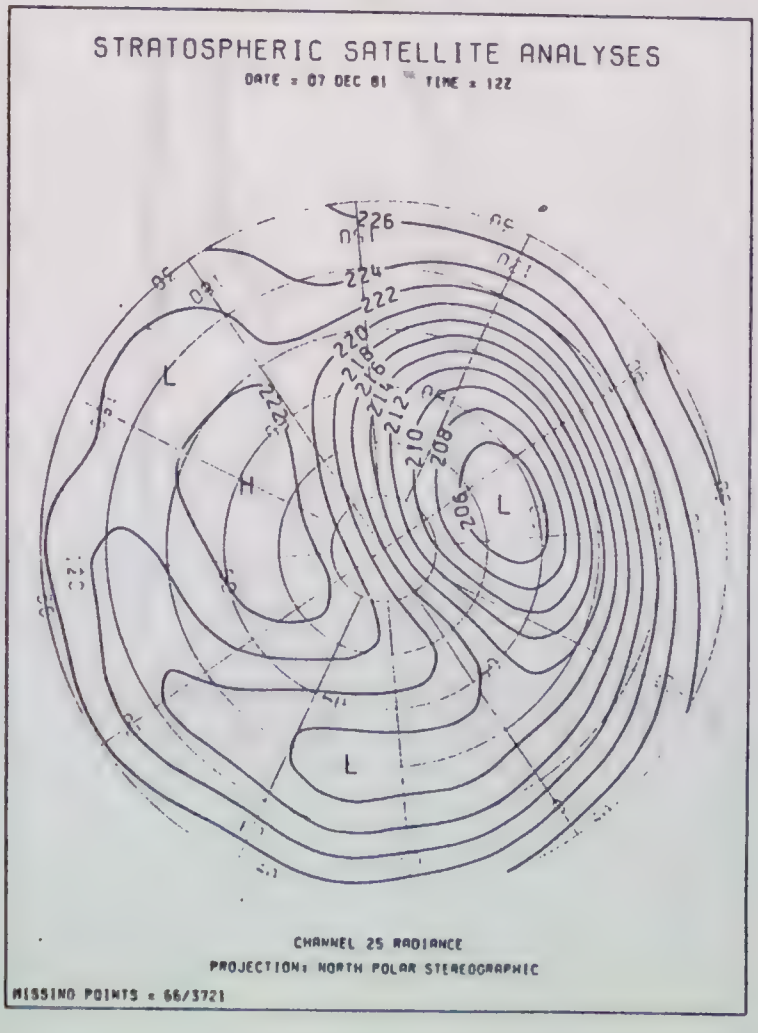

120
110
100
90
80
70
60
50
40
30
20
10
2
4
4
6
8
25
27
29
1
3
5
7
9
11
13
15
17
19
21
23
25
December 1981

Fig 6


STRATOSPHERIC SATELLITE ANALYSES
DATE = 25 NOV 81 TIME = 12Z
H
L
226
224
222
220
218
216
214
212
210
208
206
204
CHANNEL 25 RADIANCE
PROJECTION: NORTH POLAR STEREOGRAPHIC
MISSING POINTS = 156/3721

(a)


STRATOSPHERIC SATELLITE ANALYSES
DATE = 07 DEC 81 TIME = 12Z
H
L
L
L
226
224
222
220
218
216
214
212
210
208
206
CHANNEL 25 RADIANCE
PROJECTION: NORTH POLAR STEREOGRAPHIC
MISSING POINTS = 66/3721

(b)

Fig 7

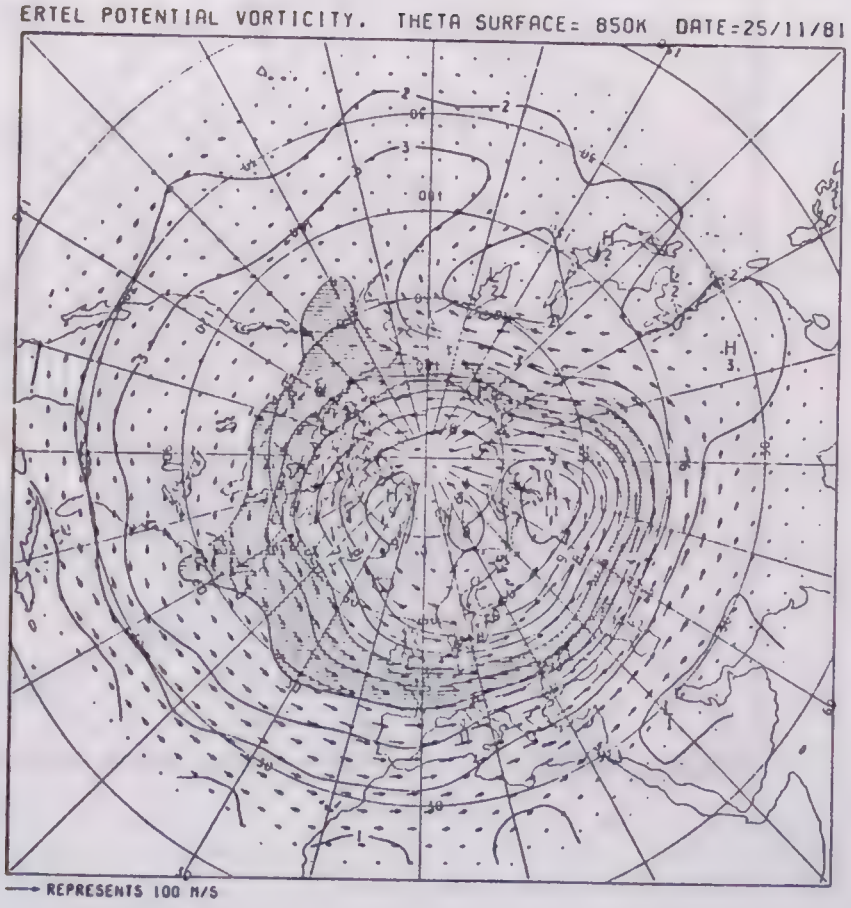

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE=25/11/81
REPRESENTS 100 M/S

Fig 8

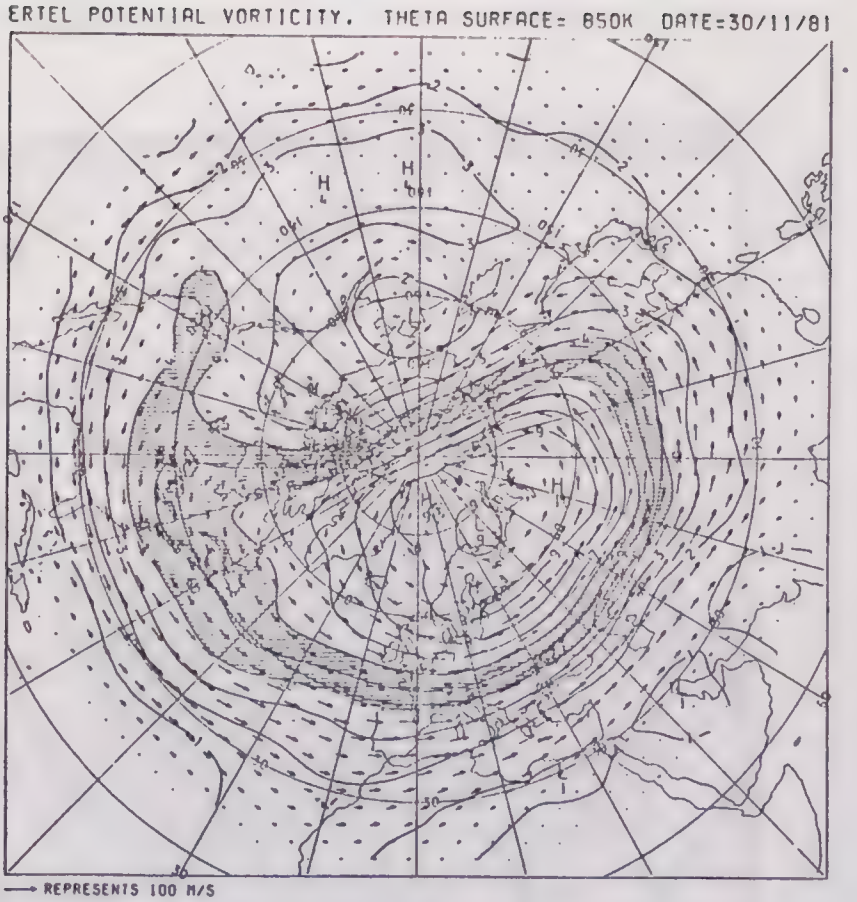

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE=30/11/81
H
H
L
REPRESENTS 100 M/S

Fig 9

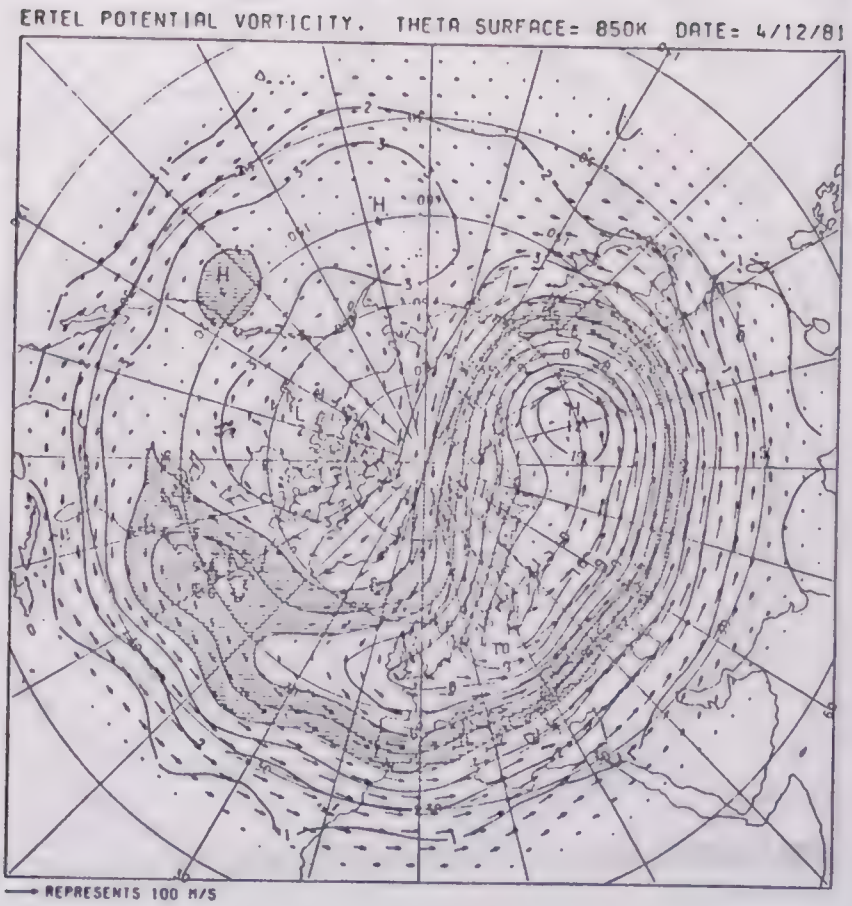

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE= 4/12/81
H
H
L
L
REPRESENTS 100 M/S

Fig 10

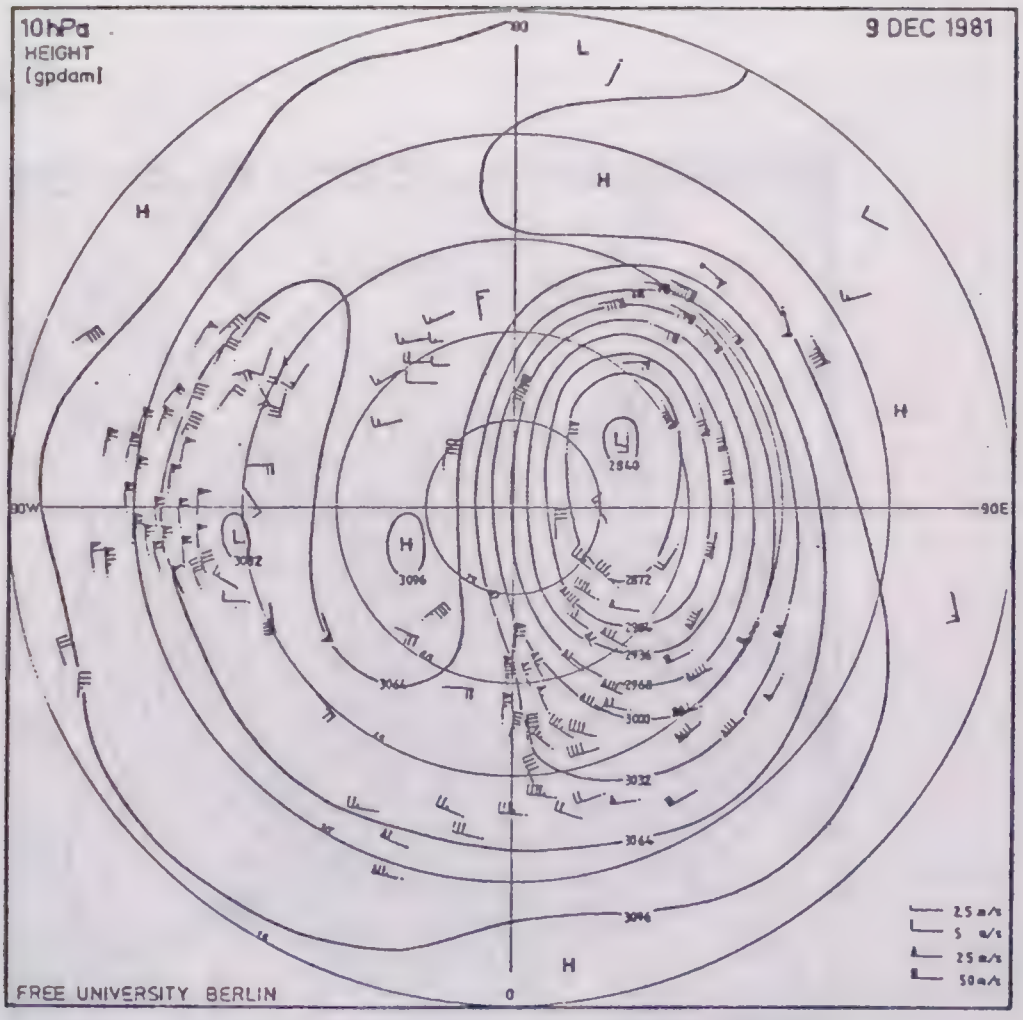

10 hPa
HEIGHT
(gpdam)
9 DEC 1981
90W
90E
0
L
H
3096
3064
3032
3000
2968
2936
2872
2840
2.5 m/s
5 m/s
25 m/s
50 m/s
FREE UNIVERSITY BERLIN

Fig 11

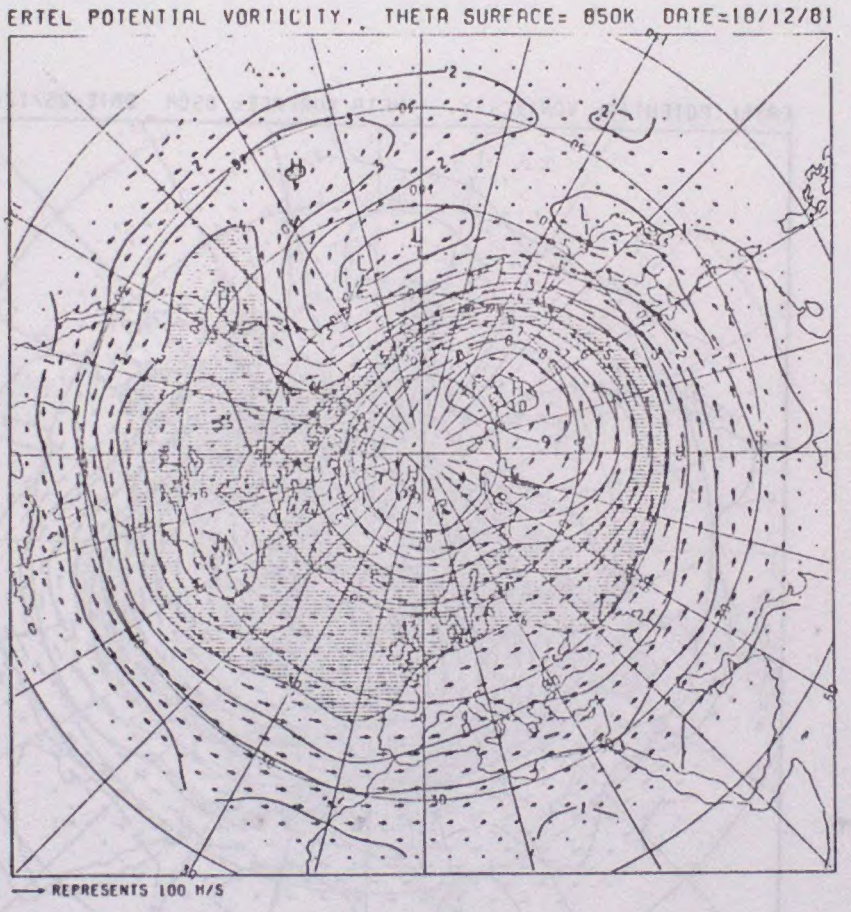

ERTEL POTENTIAL VORTICITY, THETA SURFACE= 850K DATE=18/12/81
REPRESENTS 100 M/S

Fig 12

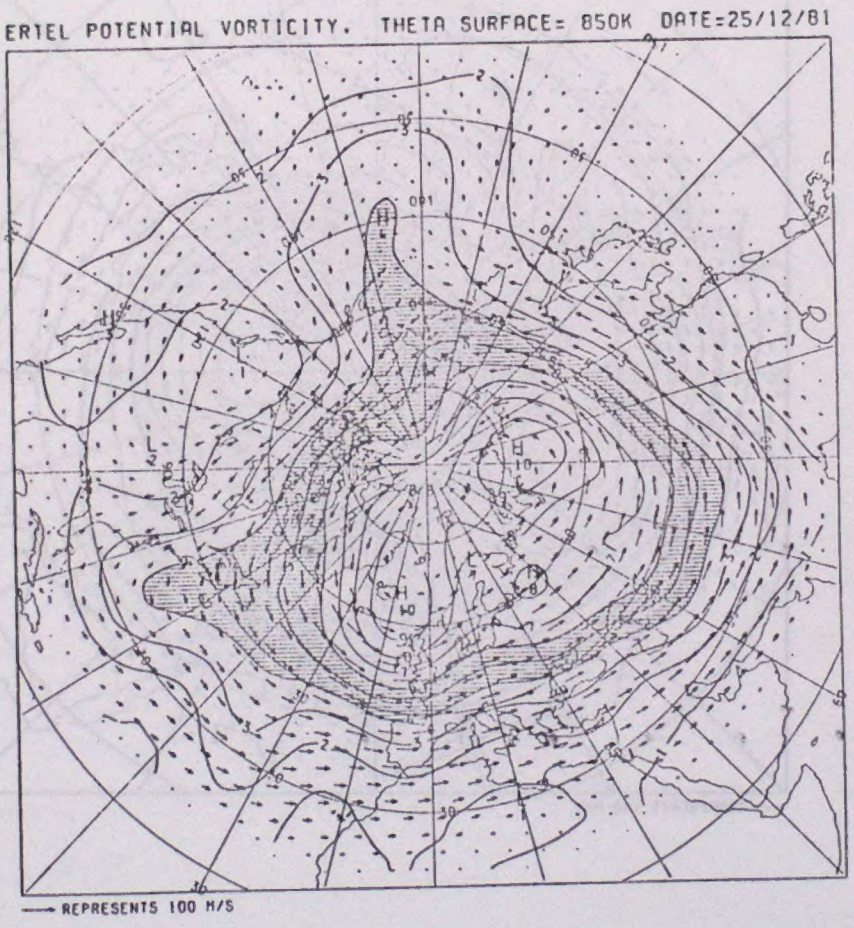

ERTEL POTENTIAL VORTICITY. THETA SURFACE= 850K DATE=25/12/81
REPRESENTS 100 M/S

Fig 13

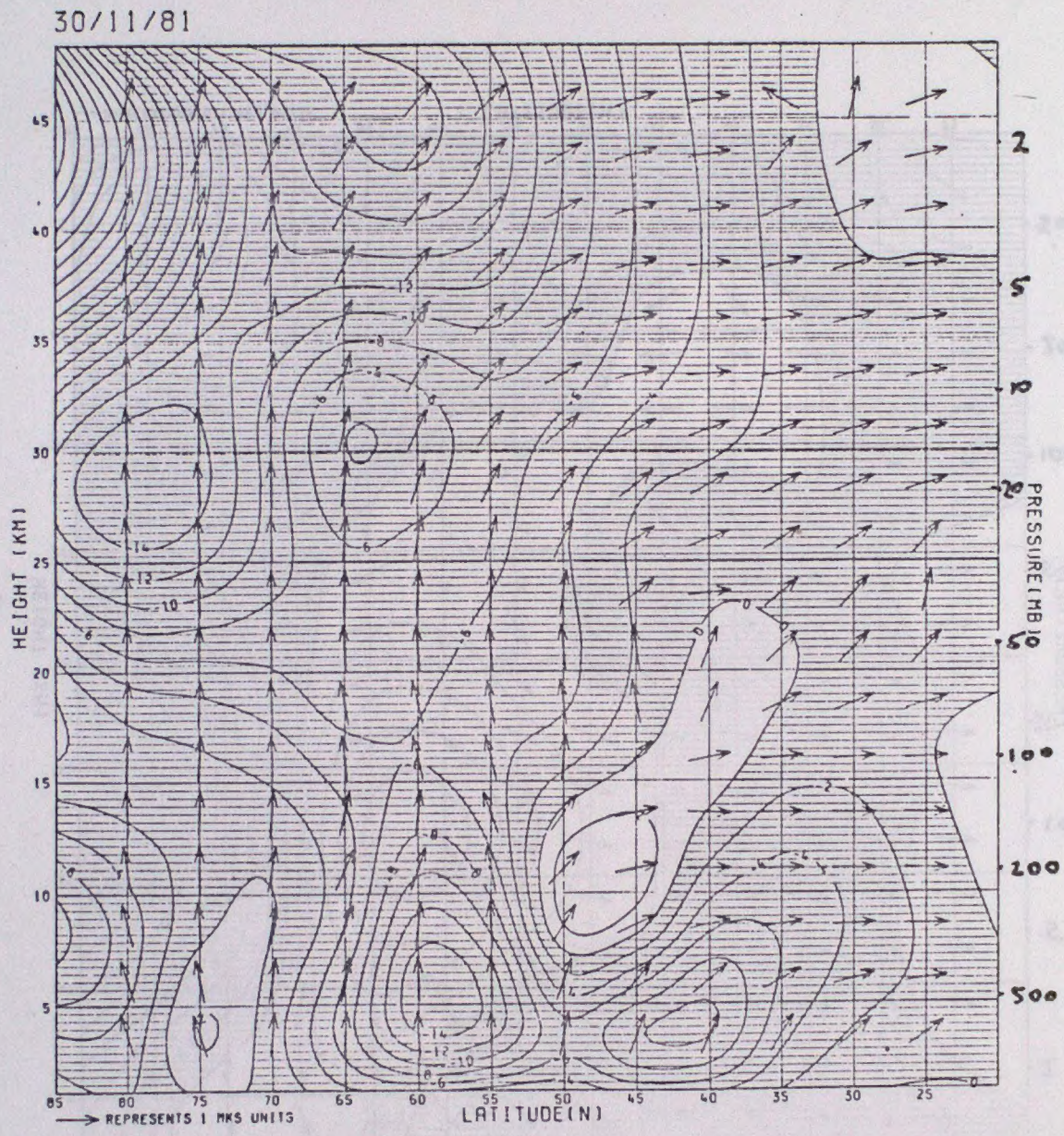
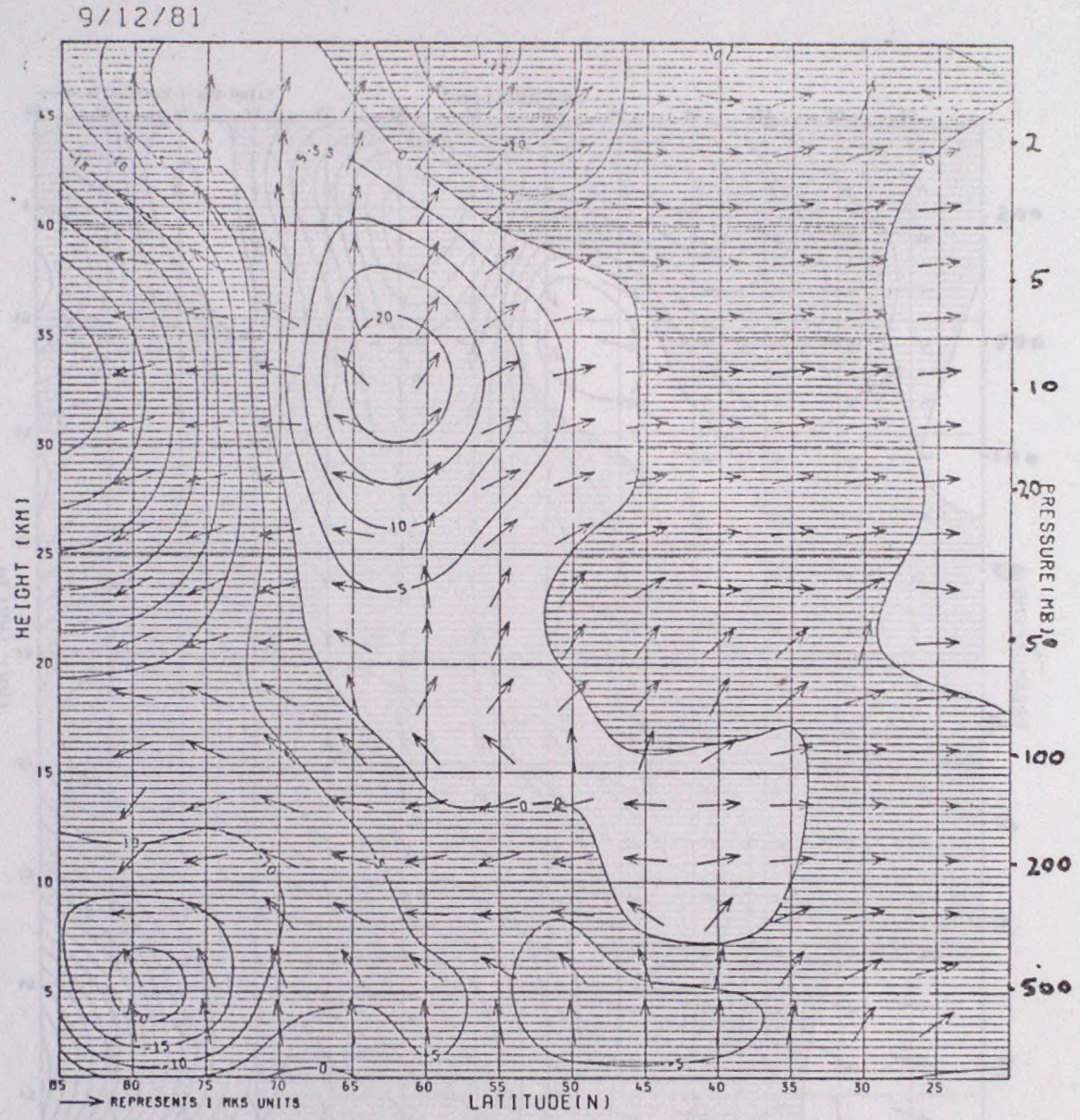

30/11/81
HEIGHT (KM)
PRESSURE (MB)
LATITUDE (N)
REPRESENTS 1 MKS UNITS

Fig 14


9/12/81
HEIGHT (KM)
PRESSURE (MB)
LATITUDE (N)
REPRESENTS 1 MKS UNITS